



A S S E S S M E N T M A N I F E S T O

A Call for the Development of
Balanced Assessment Systems



RICK STIGGINS

ASSESSMENT MANIFESTO

A Call for the Development of
Balanced Assessment Systems

Rick Stiggins

ETS Assessment Training Institute

317 SW Alder St., Suite 1200
Portland, OR 97204

800-480-3060

www.ets.org/ati

April 2008

A Call for the Development of **Balanced Assessment Systems**

A **manifesto** is a public statement of intention, belief, opinion, or policy advocating political and/or social action. Often such ardent statements run counter to conventional or dominant values and practices within the context in which they are issued. I issue this assessment manifesto because I believe that we have reached a tipping point in the evolution of our schools when we must fundamentally reevaluate, redefine, and redesign assessment's role in the development of effective schools. The work to be done is so crucial as to require urgent pedagogical, social, and political action.

This manifesto requires an immediate response for two reasons. First, society has changed the mission of its schools, requiring that assessment serve in important new ways—in ways fundamentally different from the past. Second, policymakers at all levels, school leaders, the measurement community, and school communities have been guided by a set of beliefs about what role assessment ought to play in schools—beliefs that have been so dominant that they have been unable to see and understand, let alone implement, new breakthroughs in our understanding of how to use assessment to promote student success.

Regarding the new mission, schools must no longer be places where some succeed at learning while others tumble into inevitable failure. Rather, they must become places where all students meet prespecified academic achievement standards. This change is driven by the accelerating technical and ethnic evolution of our society and the concomitant need for all students to master foundational lifelong learning proficiencies.

As a result of this change in mission, assessment practices developed to separate the successful from the unsuccessful now must become practices that support the learning of all students, helping them master those required standards. The vision of excellence in assessment presented below details the steps in that transformation. For reasons that follow, educational policymakers, school leaders, and the measurement community must oversee the completion of this transformational journey.

Some important facets of this metamorphosis of our assessment priorities already have revealed themselves. One example has been our slow but steady shift *from almost total reliance on norm-referenced interpretation of test scores to criterion-referenced interpretation of results*. This parallels the evolution of our collective thinking about the purposes of assessment. We have emerged from the era of assessing merely to rank students based on achievement to asking the key question: Who has and has not met standards?

Two other important shifts that are beginning to emerge strive to balance *formative with summative* and *large-scale with classroom* assessments. While not yet mainstream assessment priorities, these changing priorities are being driven forward by recent discovery of profound achievement gains attainable through the thoughtful application of formative classroom assessment practices (research detailed below).

These developments foreshadow more changes that are to be encouraged. Perhaps the most fundamental of these changes is the way we judge the quality of an assessment. Historically, the challenge to the measurement community has been to produce dependable scores. So research and development attention has centered on understanding how to create measurement instruments that yield quality results—that is, results that lead users to valid and reliable inferences about student achievement. Decades of increasingly sophisticated technical advances have resulted in a deep understanding of how to produce, scale, and interpret test scores that consistently and accurately reflect the intended achievement target. These have been critically important and immensely productive developments.

However, as the mission of schools has evolved toward an ever-stronger emphasis on mastery of standards and as we have come to more clearly understand how to use assessment to support student learning, it has become apparent that we must judge assessment quality based on far more than merely the dependability of the resulting scores. Quality must also include the evaluation *of the impact of those scores on the learner during the learning*. The most valid and reliable assessment in the

A Call for the Development of **Balanced Assessment Systems**

world that has the effect of causing students to give up in hopelessness cannot be regarded as productive because it does far more harm than good. Thus, quality must become a function of the instrument and its score evaluated in terms of (or considered simultaneously with) the context within and manner in which it is used. Quality control frameworks of the past have not taken into account impact on the learner. The vision of excellence in assessment framed herein places this criterion of quality at center stage.

If they are to have a productive impact on the learner, the nature of our assessment practices must continue to evolve in specific directions. For instance, the assessment results must go beyond merely providing *judgments about* to providing *rich descriptions of* student performance. In other words, if assessments are to support improvements in student learning, their results must inform students how to do better the next time. This will require communication of results that transmit sufficient understandable detail to guide the learner's actions. In such contexts, single scores or grades will not suffice.

Further, to support learning, assessments must evolve from being *isolated occasional events attached to the end of teaching to becoming an ongoing series of inter-related events that reveal changes in student learning over time*. Such evidence will reveal to the learner and the teacher not only current achievement status, but also improvements in the student's capabilities—a powerful booster of confidence and motivation.

Finally, to support learning, assessments must move beyond merely informing the instructional decisions of teachers and school leaders to informing decisions made by students, too. In the future, balanced assessment systems will need to be designed to *serve diverse purposes* by meeting the information needs of all decision makers. Historically, they have not done this.

The manifesto that follows describes a vision of the future of assessment that accounts for each of these ingredients and advocates for bold movement into that future by revealing what will happen to student achievement and school effectiveness as we proceed.

With Assessment, Purpose Is Everything

We assess for two reasons: (1) to gather evidence to inform instructional decisions and (2) to encourage students to try to learn. Both purposes must be well served for schools to be effective. The vision of excellence in assessment described herein holds that, to inform and encourage effectively, assessment systems must yield accurate information about student learning for use at several levels of decision making, and they must be used in a manner that manages the emotional dynamics of the assessment experience effectively for the learner

To assist productively in instructional decision making, regardless of the context of their use, assessments must meet three standards of quality. Each assessment must be designed to serve a specific predetermined purpose, arise from a specific predetermined definition of achievement success, and be built of high-quality ingredients so as to yield dependable results.

To productively manage the emotional dynamics of assessment experiences, we must strive to elicit a productive reaction to results from both students and their teachers. For the student, regardless of the level of achievement demonstrated, a productive reaction leaves them confident and willing to keep trying. A counterproductive response leaves the student confused, frustrated, and ready to give up in hopelessness. For the teacher, the assessment is helpful if it reveals what comes next in the learning; the assessment is counterproductive when it leaves them with no idea what to do next. We will review the conditions that must be present in the assessment environment for the results to have a productive impact—that is, to encourage learning.

The power of assessment as a school improvement tool can be tapped only by achieving a synergy between assessment quality and effective use. Historically, our attention has centered on attributes of the assessment instruments and their scores. In the future, our sense of what it means to assess well must be expanded to bring the emotional dynamics of the assessment experience into the equation.

Purpose 1:

Balanced Assessment and Productive Instructional Decision Making

Truly productive assessment systems within schools and districts serve the information needs of a wide variety of important assessment users. In other words, such systems acknowledge that a wide variety of decision makers need access to different kinds of information in different forms at different times to help students learn. If any of these users' information needs is ignored, or they are provided with misinformation due to inept assessments, ineffective decisions will result that will harm student confidence, motivation, and learning, as well as student and teacher efficacy.

For this reason, the starting place for the creation of a quality assessment for use in any particular context within any system must be a clear sense of the information needs of the intended assessment user(s)/decision maker(s). Without a sense of what kind of information will help them and, therefore, what kind of assessment must be conducted, the assessor cannot proceed productively.

Therefore, the development of a productive assessment system requires a thoughtful analysis of the full range of potential assessment users and uses within a school district. Such an analysis must begin by describing the assessment demands of the *classroom level of use*, where students, teachers, and sometimes parents make their instructional decisions. Here, assessment can serve both to support learning and to verify it. Next, the analysis must progress to the *program level* of assessment use, where teacher leaders and teams, as well as principals, curriculum personnel, and others evaluate program effectiveness with the goal of identifying particular standards that students are not mastering in order to inform decisions about how to improve instructional programs. And finally, the analysis must include the *institutional accountability and policy level* of assessment use, where resource allocation, programmatic, policy, and other decisions are made by school, district, and community leaders. These are the people to be held accountable for the quality of schools.

To devise a truly helpful assessment at any of these levels, one must begin assessment design, development, and use with sharply focused answers to the following questions:

- What are the instructional decisions to be made?
- Who will be making those decisions?
- What information will help them make good decisions?

The answers to these driving questions vary profoundly across the three levels (Stiggins, 2007).

The Classroom Level of Assessment Use

At the classroom level, the context should be one in which achievement standards have been arrayed in learning progressions that unfold within and across grade levels over time to map the learner's route to ultimate academic success. If assessment is to support learning as students ascend the progressions, then, it must serve as follows:

- Decision to be made:
What comes next in the learning?
- Made by:
Students, teachers, and sometimes parents
- Information needed:
Continuous evidence of each student's current location on the scaffolding leading to each standard

In order to know what comes next in the learning, one must know where the student is now on the learning progression. Classroom assessments must provide that information, not once a year or every few weeks, but continuously. Note that the focus of attention and decision making is on the achievement of each individual student—there is no aggregation of data across students. And, note that the question is not: Who is mastering standards? Rather, it is: How is each student doing on her or his journey up the scaffolding leading to each standard? It is never the case that, first, a student can't meet a standard and then, all at once, she can. Over time, the student ascends through progressive levels of mastery of prerequisites leading up to mastery of that standard. Ongoing classroom assessment must track

A Call for the Development of **Balanced Assessment Systems**

that progress in order to know at any point in time what comes next in the learning. Only then can it serve truly formative purposes.

The Program Level of Assessment Use

The answers to the same three driving questions are different at the program level of assessment use:

- Decision to be made:
Which standards are our students mastering or not mastering?
- Made by:
Teacher teams, teacher leaders, principals, and curriculum personnel
- Information needed:
Periodic, but frequent, evidence aggregated across classrooms revealing standards not mastered

The objective in this case is to rely on interim, benchmark, short-cycle, or common assessments every few weeks to identify aspects of instructional programs that are being effective as well as those in need of improvement. The program is in need of improvement when it fails to help students master a specific standard. This formative application of assessment will tell faculties precisely where to focus their improvement efforts and how to make those improvements in a timely manner. Note that the focus of attention in this case is the achievement standard. Users seek to identify those with which students struggle so as to bring program resources to bear more effectively on their behalf.

The Institutional Level of Assessment Use

Finally, at the institutional/policy level, the accountability question comes to the fore:

- Decision to be made:
Are enough students meeting required standards?
- Made by:
Superintendents, school boards, legislators
- Information needed:
Annual summaries of standards mastered on accountability tests

In this case, assessments serve summative, accountability purposes. It is a matter of law that schools will administer annual assessments to all students in certain grade levels, revealing the proportion of students mastering pre-established standards so as to evaluate the overall institutional impact.

Note the Differences

It is critically important that the measurement community, school leaders, and policymakers at all levels see and understand the fundamental differences in the kinds of information needed across these levels of assessment uses. For instance, note carefully the differences in the questions addressed and the information needed. The classroom level asks: How goes the journey for the student up the scaffolding to competence? The program level asks: How might our programs be improved to promote greater student success? And the institutional level asks: Are our schools as effective as they need to be? No single assessment is capable of answering all of these questions. A productive, multi-level assessment system is needed to be sure that all instructional decisions are informed and made well.

Note also the different users and uses served at the three levels. The classroom serves students as they decide whether success at learning is within reach for them, and as they decide how to approach that learning productively. It informs teachers as they track what comes next in the learning, how to promote that learning, what feedback to provide to students, and how to judge the sufficiency of each student's progress. At the program level, faculty teams can use results of common assessments to examine their professional effectiveness or the relative effectiveness of different instructional interventions in helping students master pre-set standards. At the institutional level, matters of leadership effectiveness, instructional policy, resource allocation, and other such broad program variables come under scrutiny. *Clearly, an assessment system that fails to meet the information needs of any of these important decision makers at any of these levels places students directly in harm's way due to inept instructional decision making.*

A Call for the Development of **Balanced Assessment Systems**

In other words, all parts of the system must contribute to effective schooling. If assessment isn't working effectively day to day in the classroom, then the program and institutional levels of assessment cannot pick up the slack. For example, if bad decisions are being made day to day by teachers during the learning, then there isn't an interim or annual assessment yet invented that can overcome the dire consequences for the learners. But, at the same time, equally unique and important decisions are made at program evaluation and institutional levels.

The balanced assessment systems of the future, unlike the unbalanced standardized test/accountability-driven systems of the past, must meet the information needs of all relevant assessment users.

The Critical Foundations

One structural foundation of a productive assessment system, therefore, is the framework of achievement expectations to be reflected in its component assessments. Whether those guiding achievement expectations are framed as state standards, local standards, a teacher's classroom standards, or the local curriculum designed to take students over time to those standards, certain criteria must be satisfied. For instance, our academic standards must be:

- Centered on the truly important learnings of the field of study
- Clearly and completely integrated into learning progressions within and across grades (Heritage, no date)
- Precisely defined such that qualified educators consistently interpret them to mean the same thing
- Within developmental reach of the students who are to master them
- Manageable in number given the resources available to teach and learn them
- Thoroughly mastered by those teachers charged with helping students master them

If these criteria are not met, then both quality assessment and effective instruction will remain beyond reach. So the starting place for the development of balanced

assessment systems is the verification of the quality of the learning expectations upon which it will rest. Until each local set of standards is in order, further consideration of assessment quality and use will be pointless.

In this same spirit, a second foundation of an effective, balanced assessment is a commitment to the development and implementation of standards-based schools. Faculty must understand what it means to design and offer standards-based instruction, and they must be committed to a mission of maximizing the success of each student in mastering the standards in question. Without these components, focus will be missing, as will the willingness to invest in extensive student success.

A third foundation of a productive assessment system is quality assessment. To yield dependable results, regardless of the context of their use, assessments must meet these standards of quality: They must be designed to serve a specific predetermined purpose, arise from a specific predetermined definition of achievement success, be designed specifically to fit into each particular purpose and target context, and communicate their results effectively (Stiggins, Arter, Chappuis, and Chappuis, 2006). When an assessment is of high quality, it is sensitive enough to detect and accurately reflect changes in student achievement that evolve over time. Classroom, interim benchmark, or state assessments not accurate or sensitive enough to detect such changes will not contribute to productive assessment systems or school improvement (Popham, 2008).

A Note of Caution

In states and districts where these foundations are in place, excellence in assessment is within reach. However, caution is warranted here. The fact is, they may not be in place in some states or districts. Judgments of the inadequacy of state standards are quite common, as are concerns about local communities' and educators' understanding of and commitment to standards-based schools.

But more importantly, it remains the case that virtually all teachers completed pre-service preparation programs devoid of the kind and quality of assessment training

A Call for the Development of **Balanced Assessment Systems**

needed to fulfill classroom assessment responsibilities described above. And, lest we believe that they can turn to their principals for help, we must also face the fact that assessment training remains nearly nonexistent in university-based leadership preparation programs nationwide. This should not lead to the automatic inference that local educators in any particular district lack the required assessment literacy. But the opportunities to learn sound assessment practices in an in-service context also remain very, very rare.

These foundations are essential to the development of productive, balanced assessment systems, and it is essential that they receive careful attention to be sure they are in place before moving forward in school improvement endeavors.

Purpose 2:

Productive Assessment Dynamics and Student Success

No one will question the need for accurate evidence effectively used to inform all relevant instructional decisions as described above. In moving forward on that agenda, our historic assessment paradigm endures, requiring only some refinements on its application. In the future, all educators must be assessment literate, not just the professional test developers.

However, there is another part of my vision of excellence in assessment that represents a much more profound shift in paradigm. It centers on the emotional dynamics of the assessment experience from the student's point of view.

When the mission of schools was to rank students (instead of also assuring that all students meet pre-established standards as now), the amount of time available to learn was fixed: one year per grade. The amount learned by the end of that time varied—some of us learned a great deal, some very little. Able learners built on past success to grow rapidly. However, students who failed to master the early prerequisites within the allotted

time also failed to learn that which followed. After 13 years of cumulative treatment in this manner we were, in effect, spread along an achievement continuum that labeled each student's rank in class upon graduation.

The emotional dynamics of this process were clear and purposeful. From the very earliest grades, some students rode winning streaks to the top. Right from the start, they scored high on assessments and were assigned high grades. The emotional effect of this was that they came to see themselves as capable learners—they became increasingly confident in school. That gave them the emotional strength to risk striving for more success because in their minds success was within reach if they tried. Note that the trigger for the decisions they made about their own learning was their interpretation of their own assessment results.

But other students scored very low on tests right from the beginning and so they were assigned failing grades. This caused them to doubt their own capabilities as learners from the outset. Their loss of confidence deprived them of the emotional reserves to continue to risk trying. Chronic failure was hard to hide and became embarrassing—better not to try. As their motivation waned, of course, their achievement suffered. Notice again how the learners' own interpretation of assessment results influenced their confidence and willingness to strive on.

In these schools, if some students worked hard and learned a great deal, that was a positive result, as they would finish high in the rank order. And, if some students gave up in the face of what they believed to be inevitable failure, that was an acceptable result for the institution too, because they would occupy places very low in the rank order. The greater the spread of achievement from top to bottom, the more dependable would be the rank order. Mission accomplished. This is why, if a student gave up and stopped trying (even dropped out of school), it was regarded as that student's problem, not the teacher's or school's problem. The school's responsibility was to provide the opportunity to learn. If students didn't take advantage of the opportunity, that was not the system's responsibility.

A Call for the Development of **Balanced Assessment Systems**

The important lesson we must learn is that the student's emotional reaction to any set of assessment results, whether high, mid-range, or low, will determine what the student thinks, feels, and does in response to those results. They can respond in either of two ways to any set of assessment results, one productive and the other not. The productive reaction leaves students saying, "I understand these results. I know what to do next to learn more. I can handle this. I choose to keep trying." The counter-productive response leaves students saying, "I don't know what these results mean for me." Or, "I have no idea what to do next." Or, "I'm too dense, slow, and stupid to learn this. I quit."

When those who give up in hopelessness are those who also have yet to meet standards, and if educators are to be held accountable by society for all students meeting standards, we have a serious problem. These students will stop trying and will neither master essential foundational reading, writing, and math problem-solving proficiencies nor become lifelong learners.

If society wants all students to meet standards, then all students must believe they can meet those standards; they all must be confident enough to be willing to take the risk of trying. Any other emotional state for any student is unacceptable. We simply can no longer have students who have yet to meet standards losing faith in themselves and giving up in futility.

In other words, assessment practices that permit—even encourage—some students to give up on learning must be replaced by those that engender hope and sustained effort for all students. If all students are to meet standards, the emotional environment surrounding the experience of being evaluated must change for all, but especially for perennial low achievers. The driving emotional force of intimidation and fear triggered by the prospect of being held accountable now must be replaced by the driving emotions of optimism and persistence triggered by the belief that "I can succeed at learning if I try." If all students are to succeed, they must have continuous access to credible evidence of their own academic *success* at mastering prescribed achievement standards.

I believe that the importance of this change in our assessment paradigm cannot be overstated. Over the decades, school improvement experts have made the mistake of believing that the adults in the system are the most important assessment user or data-based instructional decision makers—that is, we have believed that, as the adults make better instructional decisions, schools will become more effective. Clearly parents, teachers, school leaders, and policy makers make crucial decisions that influence the quality of schools, and the more data-based those decisions are, the better. But this perspective overlooks the reality that students may be even more important data-based instructional decision makers than the adults.

Consider, for example, that students are constantly asking themselves, "Can I get this, or is it just too hard for me? Is the learning worth the energy I must expend to attain it? Is the learning worth the risk of public failure?" We must understand that, if students come down on the wrong side of these crucial decisions and thus stop trying, it doesn't matter what the adults around them decide. The learning stops. In this sense, students can render their teachers' instructional decisions null and void. They have it within their power to make the adults ineffective and to prevent them from doing anything about it. If a student decides that the learning is beyond reach for her or him or that the risk of public failure is too great and too embarrassing, then regardless of what we adults do, the learning stops.

So the essential question for teachers and school leaders is: What can we do to help students answer the above questions in productive ways that keep them believing that success is within reach for them if they keep trying? In fact, we know how to do this through effective classroom assessment, and we know what will happen to student achievement when we put effective classroom assessment practices in place.

Using Assessment for Learning

Classroom assessment *for* student learning, as defined herein, turns the classroom assessment process and its results into an instructional intervention designed

A Call for the Development of **Balanced Assessment Systems**

to increase, not merely monitor, student confidence, motivation, and learning (Stiggins, Arter, Chappuis and Chappuis, 2004; William, 2007). Research evidence gathered in studies conducted literally around the world over the past two decades shows that the consistent application of principles of assessment *for learning* can give rise to profound gains in student achievement, especially for perennial low achievers.

One unique feature of the assessment *for learning* process is that it acknowledges the critical importance of the instructional decisions made by students working in collaboration with their teachers. In this case, assessment provides students with information about their own achievement improvement and status when they need it. In that context, students become consumers of assessment information too, using evidence both to see their current successes and to understand what comes next for them. If done well, it elicits a productive response from learners every time.

Another important feature of assessment *for learning* is its reliance on repeated self-assessments, each of which instructs the learner on how to improve performance on the next one. This kind of continuous descriptive feedback provided strategically in amounts that students can address effectively and that builds progressively over time helps them continue to believe that success is within reach if they keep trying.

Still another unique feature of this process is its reliance on carefully drawn learning progressions or curriculum maps written in teacher-, student- and family-friendly versions so that the trajectory (i.e., what has been learned and what comes next) is clear to all throughout the learning. This, like the descriptive feedback above, leads directly to our second reason for assessing: If we assess to motivate students to try, assessment *for learning* enables students by helping them watch themselves grow—by causing them to believe that success is within reach if they keep trying.

The psychological underpinnings of student motivation and learning success are directly relevant here. Our aspiration is to give each student a strong sense of control over her or his own academic well-being. Albert

Bandura (1994) refers to this sense as “self-efficacy.” In the paragraphs that follow, he describes this continuum as a psychological construct. However, if the reader will think of this continuum in terms of the student’s sense of control over learning success (academic self-efficacy, if you will), it will become clear that the consistent application of principles of assessment *for learning* can move students boldly toward the productive end:

A strong sense of efficacy enhances human accomplishment and personal well-being in many ways. People with high assurance in their capabilities approach difficult tasks as challenges to be mastered rather than as threats to be avoided. Such an efficacious outlook fosters intrinsic interest and deep engrossment in activities. They set themselves challenging goals and maintain strong commitment to them. They heighten and sustain their efforts in the face of failure. They quickly recover their sense of efficacy after failures or setbacks. They attribute failure to insufficient effort or deficient knowledge and skills which are acquirable. They approach threatening situations with assurance that they can exercise control over them. Such an efficacious outlook produces personal accomplishments, reduces stress and lowers vulnerability...

In contrast, people who doubt their capabilities shy away from difficult tasks which they view as personal threats. They have low aspirations and weak commitment to the goals they choose to pursue. When faced with difficult tasks, they dwell on their personal deficiencies, on the obstacles they will encounter, and all kinds of adverse outcomes rather than concentrate on how to perform successfully. They slacken their efforts and give up quickly in the face of difficulties. They are slow to recover their sense of efficacy following failure or setbacks. Because they view insufficient performance as deficient aptitude it does not require much failure for them to lose faith in their capabilities. (p. 71)

We help students build a strong sense of academic self-efficacy when we help them understand that their role in the assessment environment is to strive to understand what success looks like and when we show them how

A Call for the Development of **Balanced Assessment Systems**

to use each assessment to determine how to do better the next time. Assessments become far more than merely one-time events attached to the end of the teaching. They become part of the learning process by keeping students posted on their progress and confident enough to continue striving. Students become partners in the self-assessment process during the learning by, for example, collaborating with their teachers in the creation and use of assessments like those they will be held accountable for later. This reveals to them the secrets to their own learning success while they are still learning. They can become partners in the accumulation of growth portfolios that reveal to them, their teachers, and their families changes in their own achievement as it is happening. This builds confidence that ultimate success is always within reach. Finally, students can become partners in communicating about their own learning success as they rely on concrete evidence from their portfolios presented in student-led conferences to inform their families of their learning.

When assessment *for* learning practices like these play out as a matter of routine in classrooms, as mentioned previously, evidence gathered from dozens of studies conducted around the world consistently reveals a half to a full standard deviation gain in student achievement attributable to the careful management of the classroom assessment process, with the largest gains accruing for struggling learners. (Black and Wiliam, 1998; Hattie and Timperley, 2007).

Assessment Manifesto: A Total Assessment Solution

We assess for two reasons: to inform decisions and to motivate students. As educators, we have maintained the faith that both uses of assessment can enhance the quality of schools. Recently, that faith has been bolstered with research and reason. We understand far more today than ever before about how to use assessment productively. We must replace grossly out-of-balance assessment systems of the past with those that honor the information needs of all assessment users—systems that both support and verify learning from the classroom to the boardroom.

To attain long-missing and much-needed balance, we must implement classroom assessment practices that rely on an ongoing array of quality assessments used strategically in ways that keep students believing in themselves. If we do so, we can realize profound gains in achievement. In other words, it is time to replace the intimidation of accountability as our prime motivator with the promise of academic success for all learners as that motivational force. It is not that intimidation is universally ineffective, but it only motivates those who have hope of success. *Unfortunately, true hopelessness always trumps intimidation when it comes to learning.* Effective classroom assessment can and must serve to promote hope in all students.

There are several reasons why we have had difficulty understanding the need for balance in our assessment systems at the highest levels of school policy and practice. Historically, educational leaders and teachers have not been given the opportunity to learn about sound classroom assessment practices. Further, over the decades, the measurement community has narrowed its role to one of maximizing the efficiency and accuracy of high-stakes testing while paying virtually no attention to assessment as it plays out for teachers or learners day to day in the classroom. The business community has believed that we get better schools by comparing schools based on annual test scores and rewarding or punishing them. They fail to understand the negative impact on schools that continuously lose in this competition — that is, students' and teachers' loss of efficacy in perennially struggling schools. Politicians at all levels have believed that, if a little intimidation doesn't work, a lot of intimidation will, and assessment has represented the way to intensify anxiety. They too misunderstand the implications of such testing priorities on struggling schools and, more importantly, on struggling learners.

Yet, research and development efforts conducted over the past two decades leave us positioned to transform assessment systems in productive ways. We have come to understand the power of assessment balance. For the first time in the evolving history of assessment in America,

A Call for the Development of **Balanced Assessment Systems**

professional development programs in classroom assessment for teachers and balanced assessment for school leaders make it possible to help local educators develop assessment literacy. Such literacy is needed to design and build totally integrated assessment systems with all parts working together in the service of student success. While virtually all state licensing standards require competence in assessment, typically neither pre-service nor in-service teacher or administrator training programs include this kind of training (Crooks, 1989; Black and Wiliam, 1998; Stiggins, 1999; Shepard et al, 2005). The tools are now readily available for faculties of education to change this—to teach sound assessment practices by modeling them.

The current state of affairs is clear: We know what teachers and administrators need to know and understand to assess effectively day to day or year to year. We can

provide them with the assessment tools and technologies needed to assess effectively. Further, it is crystal clear what will happen to student learning if educators properly perform assessments. And we know how to deliver the proper assessment competencies and tools into the hands of all key users with efficient and effective professional development. The only unanswered question is: *Will practitioners and policymakers be given the opportunity to learn to assess productively?* Historically, the answer has been an unequivocal, "No, they will not." As a result, the immense potential of assessment to support student learning has gone untapped—indeed, unnoticed at the highest levels of policy making. It need not be so. We have in hand a new vision of excellence in assessment that will tap the wellspring of confidence, motivation, and learning potential that resides within every student. The time has come to embrace it.

Thanks to Steve Chappuis, Jan Chappuis, Jim Popham and Lorrie Shephard for providing feedback on early drafts.

References

- Bandura, A. (1994) Self-efficacy. In V.S. Ramachaudran (ed.) *Encyclopedia of human behavior*. Volume 4, 71-81. New York: Academic Press.
- Black, P. and Wiliam, D. (1998). Assessment and classroom learning. *Educational Assessment: Principles, Policy and Practice*, 5(1), 7-74. Also summarized in an article entitled, Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- Crooks, T.J. (1988). The impact of classroom evaluation on students. *Review of Educational Research*, 58(4), 438-481.
- Hattie, J. and Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Heritage, M. (no date). Learning progressions: Supporting instruction and formative assessment. Washington DC: A paper published by the Council of Chief School Officers. 30 pp.
- Popham, W.J. (2008). *Transformative Assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Shepard, L., Hammerness, K., Darling-Hammond, L., Rust, F., Snowden, J. B., Gordon, E., Gutierrez, C., and Pacheco, J. (2005). Chapter 8 in Darling-Hammond, L. and Bransford, J. (Eds) *Preparing teachers for a changing world: What teachers should know and be able to do*. San Francisco, CA: Jossey-Bass (note particularly pp. 282-284).
- Stiggins, R.J. (1999). Evaluating classroom assessment training in teacher education. *Educational Measurement: Issues and Practice*, 18 (1), 23-27.
- Stiggins, R.J., Arter, J., Chappuis, J. and Chappuis, S. (2006). *Classroom Assessment for Student Learning: Doing It Right—Using It Well*. Portland OR: ETS Assessment Training Institute.
- Stiggins, R.J. (2007). Assessment for learning: A key to student motivation and learning. *Phi Delta Kappa EDGE*, 2(2), 19pp.
- William, D. (2007). Changing classroom practice. *Educational Leadership*. 65(4), 36-42.

ASSESSMENT MANIFESTO

A Call for the Development of Balanced Assessment Systems

ETS Assessment Training Institute
317 SW Alder St., Suite 1200
Portland, OR 97204

800-480-3060

www.ets.org/ati



Listening. Learning. Leading.®

www.ets.org